2. Basic concepts of codes and the generic coding scheme

Coding Technology

Problem 1

A code has codewords

$$10100, 01111, 11110, 00000.$$

(a) Calculate the n and k parameters of the code.

(b) What is the minimal Hamming distance between codewords?

(c) How many errors can the code detect? How many errors can the code correct?

Problem 1

Solution.

(a) The length of the codewords is $n = 5$, and the number of codewords is $2^k = 4$ (one for each message vector of length k), so $k = 2$. This is a $C(5, 2)$ code.

Problem 1

Solution.

(a) The length of the codewords is $n = 5$, and the number of codewords is $2^k = 4$ (one for each message vector of length k), so $k = 2$. This is a $C(5, 2)$ code.

(b) Using pairwise comparison, the minimal Hamming distance is

$$d_{\min} = 2.$$

Problem 1

Solution.

(a) The length of the codewords is $n = 5$, and the number of codewords is $2^k = 4$ (one for each message vector of length k), so $k = 2$. This is a $C(5, 2)$ code.

(b) Using pairwise comparison, the minimal Hamming distance is

$$d_{\text{min}} = 2.$$

(c) A code with $d_{\text{min}} = 2$ can detect

$$d_{\text{min}} - 1 = 1$$

errors and correct

$$\left\lfloor \frac{d_{\text{min}} - 1}{2} \right\rfloor = 0$$

errors.

Problem 2

(a) Design a C(5,2) code with maximal d_{min}.

(b) Implement the code with Look-Up-Tables (LUT).

(c) Determine the error correction and error detecting capabilities of the code.

Problem 2

(a) Design a C(5,2) code with maximal $d_{\min}$.

(b) Implement the code with Look-Up-Tables (LUT).

(c) Determine the error correction and error detecting capabilities of the code.

Solution.

(a) There are 32 binary vectors of length $n = 5$, and we have to choose $2^k = 2^2 = 4$ of them. We need to check their minimal pairwise Hamming distance, and choose 4 vectors where the minimal Hamming distance is as large as possible.

Problem 2

(a) For the code in the previous problem, $d_{\text{min}} = 2$. Try to get $d_{\text{min}} = 3$.

- The first codeword could be (00000).

Problem 2

(a) For the code in the previous problem, $d_{\text{min}} = 2$. Try to get $d_{\text{min}} = 3$.

- The first codeword could be (00000).
- The weight of every other codeword must be 3 or more. (00111) is a natural choice.

Problem 2

(a) For the code in the previous problem, $d_{\text{min}} = 2$. Try to get $d_{\text{min}} = 3$.

- The first codeword could be (00000).
- The weight of every other codeword must be 3 or more. (00111) is a natural choice.
- The next codeword also needs to have weight 3 or more, but it also has to differ from the previous codeword in at least 3 digits. (11100) is a suitable choice.

Problem 2

(a) For the code in the previous problem, $d_{\text{min}} = 2$. Try to get $d_{\text{min}} = 3$.

- The first codeword could be (00000).
- The weight of every other codeword must be 3 or more. (00111) is a natural choice.
- The next codeword also needs to have weight 3 or more, but it also has to differ from the previous codeword in at least 3 digits. (11100) is a suitable choice.
- (11111) is not suitable for the final codeword, because its Hamming distance from the previous two codewords is just 2. But if we change the third bit to 0, it works: (11011).

For this code, $d_{\text{min}} = 3$.

Problem 2

(a) For the code in the previous problem, $d_{\text{min}} = 2$. Try to get $d_{\text{min}} = 3$.

- The first codeword could be (00000).
- The weight of every other codeword must be 3 or more. (00111) is a natural choice.
- The next codeword also needs to have weight 3 or more, but it also has to differ from the previous codeword in at least 3 digits. (11100) is a suitable choice.
- (11111) is not suitable for the final codeword, because its Hamming distance from the previous two codewords is just 2. But if we change the third bit to 0, it works: (11011).

For this code, $d_{\text{min}} = 3$.

4 codewords of length 5 with $d_{\text{min}} \geq 4$ is not possible. (Even 3 codewords of length 5 with $d_{\text{min}} \geq 4$ is not possible. Why?)

Problem 2

(b) If we assign the messages to the codewords according to the following list, then the lookup-table (LUT) is the same assignment in reverse:

$$\begin{array}{l} 00 \to 00000 \\ 01 \to 11100 \\ 10 \to 00111 \\ 11 \to 11011 \end{array} \quad \to$$

c'	u'
00000	00
11100	01
00111	10
11011	11

The LUT above includes only the codewords, other received vectors are decoded to the codeword with minimal Hamming-distance.

(In case of $d = 2$, the received vector may have minimal Hamming-distance to multiple codewords. In such a situation, we may choose any of the codewords with minimal Hamming-distance for decoding.)

Problem 2

(b) The full $c' \to u'$ assignment for all possible received vectors is as follows:

c'	u'
00000, 00001, 00010, 00100, 01000, 10000, 01001, 10001, 10010	00
11100,11101,11110,11000,10100,01100, 01101, 01110	01
00111,00110, 00101, 00011, 01111, 10111, 10110, 10101	10
11011, 1110, 11001, 11111, 10011, 01011, 01010	11

This table is significantly larger than the LUT, but it is not necessary to compute $\{c : \min d(c, c')\}$.

Problem 2

(b) The full $c' \to u'$ assignment for all possible received vectors is as follows:

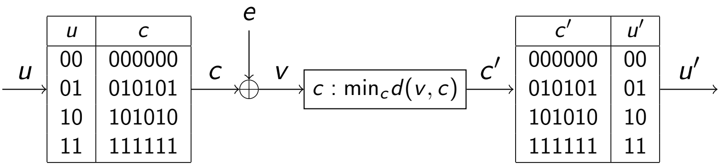
c'	u'
00000, 00001, 00010, 00100, 01000, 10000, 01001, 10001, 10010	00
11100,11101,11110,11000,10100,01100, 01101, 01110	01
00111,00110, 00101, 00011, 01111, 10111, 10110, 10101	10
11011, 1110, 11001, 11111, 10011, 01011, 01010	11

This table is significantly larger than the LUT, but it is not necessary to compute $\{c : \min d(c, c')\}$.

(c) Error detection: $d_{\min} - 1 = 2$.
Error correction: $\lfloor \frac{d_{\min}-1}{2} \rfloor = 1$.

Problem 3

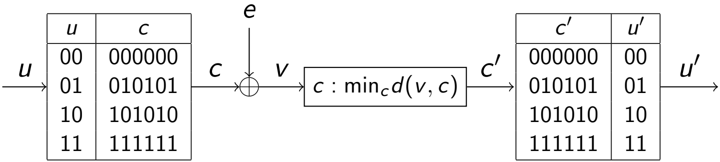
We have the following coding scheme:



For $u = (11)$ and $e = (001000)$, determine the vectors c, v, c', u'.

Problem 3

We have the following coding scheme:



u	c
00	000000
01	010101
10	101010
11	111111

c'	u'
000000	00
010101	01
101010	10
111111	11

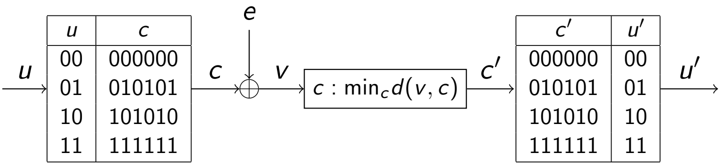
For $u = (11)$ and $e = (001000)$, determine the vectors c, v, c', u'.

Solution.

$$\begin{aligned} c &= (111111) \\ v &= c + e = (111111) + (001000) = (110111), \\ c' &= \{c : \min_{c} d(v, c)\} = (111111) \\ u' &= (11) \end{aligned}$$

Problem 4

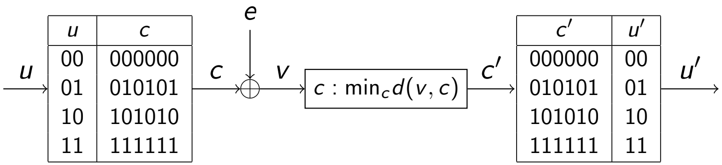
Use the same coding scheme:



for $u = (01)$ and $e = (001011)$ to determine the vectors c, v, c', u'.

Problem 4

Use the same coding scheme:



for $u = (01)$ and $e = (001011)$ to determine the vectors c, v, c', u'.

Solution.

$$
\begin{aligned}
c &= (010101) \\
v &= c + e = (010101) + (001011) = (011110), \\
c' &= \{c : \min_c d(v, c)\} = (111111) \\
u' &= (11)
\end{aligned}
$$

Problem 5

For each of the following sets of codewords, give the appropriate (n, k, d) designation, where n is number of bits in each codeword, k is the number of message bits transmitted by each codeword and $d = d_{\text{min}}$ is the minimum Hamming distance between codewords. Also give the code rate.

(a) $\{111, 100, 010, 001\}$

(b) $\{00000, 01111, 10100, 11011\}$

(c) $\{00000\}$

Problem 5

Solution.

(a) $\{111, 100, 010, 001\}$

- $n = 3$ (the length of the codewords);
- $k = 2$ (the number of codewords is $4 = 2^k$);
- $d = d_{\min} = 2$ (from pairwise comparison).
- the code rate is $k/n = 2/3$.

Problem 5

Solution.

(a) $\{111, 100, 010, 001\}$

- $n = 3$ (the length of the codewords);
- $k = 2$ (the number of codewords is $4 = 2^k$);
- $d = d_{\min} = 2$ (from pairwise comparison).
- the code rate is $k/n = 2/3$.

(b) $\{00000, 01111, 10100, 11011\}$

$n = 5, k = 2, d = 2$, the code rate is $2/5$.

Problem 5

Solution.

(a) $\{111, 100, 010, 001\}$

- $n = 3$ (the length of the codewords);
- $k = 2$ (the number of codewords is $4 = 2^k$);
- $d = d_{\min} = 2$ (from pairwise comparison).
- the code rate is $k/n = 2/3$.

(b) $\{00000, 01111, 10100, 11011\}$

$n = 5, k = 2, d = 2$, the code rate is $2/5$.

(c) $\{00000\}$

A bit of a trick question. $n = 5, k = 0$, d is undefined. The code rate is 0.

With only one codeword, no useful information can be transmitted since the receiver knows in advance what the codeword is going to be.

Problem 6

The Registrar has asked for an encoding of classes according to year ("Freshman", "Sophomore", "Junior", "Senior") that allows single error correction. Give an appropriate 5-bit binary encoding for each of the four years.

Problem 6

The Registrar has asked for an encoding of classes according to year ("Freshman", "Sophomore", "Junior", "Senior") that allows single error correction. Give an appropriate 5-bit binary encoding for each of the four years.

Solution. We need a $C(5, 2, 3)$ block code. We have seen one like that in Problem 2, that code is suitable for this problem too: $\{(00000), (00111), (11100), (11011)\}$.

Problem 7

Pairwise Communications has developed a block code with three data bits (D_1, D_2, D_3) and three parity bits (P_1, P_2, P_3):

$$P_1 = D_1 + D_2, \quad P_2 = D_2 + D_3, \quad P_3 = D_3 + D_1.$$

(a) Calculate (n, k) for this code.

(b) What are the codewords?

(c) What is the minimum Hamming distance of the code?

Problem 7

Solution.

(a) $n = 6, k = 3.$

Problem 7

Solution.

(a) $n = 6, k = 3.$

(b) The 8 possible codewords:

$$(000000), \quad (001011), \quad (010110), \quad (011101),$$
$$(100101), \quad (101110), \quad (110011), \quad (111000).$$

Problem 7

Solution.

(a) $n = 6, k = 3$.

(b) The 8 possible codewords:

$$\begin{array}{cccc} (000000), & (001011), & (010110), & (011101), \\ (100101), & (101110), & (110011), & (111000). \end{array}$$

(c) By pairwise inspection, the minimum Hamming distance is $d_{\text{min}} = 3$. OR (jumping ahead): this code is a linear code, and for linear codes,

$$d_{\text{min}} = \min_{c \neq (00...0)} w(c) = 3.$$

Problem 8

The receiver computes three syndrome bits E_1, E_2 and E_3 from the (possibly corrupted) received data and parity bits:

$$E_1 = D_1 + D_2 + P_1, \quad E_2 = D_2 + D_3 + P_2, \quad E_3 = D_3 + D_1 + P_3.$$

The receiver performs maximum likelihood decoding using the syndrome bits. Determine the result of the maximum-likelihood decoder from among the following:

- no errors, or
- single error in a specific bit (state which one), or
- multiple errors,

for each of the following combination of syndrome bits:

$$E_1 E_2 E_3 = 000, \qquad E_1 E_2 E_3 = 010,$$
$$E_1 E_2 E_3 = 101, \qquad E_1 E_2 E_3 = 111.$$

Problem 8

Solution. Main points to consider:

- no errors result in 0 for all of E_1, E_2 and E_3.
- if there is only one error, and it is from among D_1, D_2 or D_3, then two of E_1, E_2 and E_3 will be 1's.
- if there is only one error, and it is from among P_1, P_2 or P_3, then one of E_1, E_2 and E_3 will be 1's.
- The ML decoder will pick the result with the fewest errors.

Problem 8

Solution. Main points to consider:

- no errors result in 0 for all of E_1, E_2 and E_3.
- if there is only one error, and it is from among D_1, D_2 or D_3, then two of E_1, E_2 and E_3 will be 1's.
- if there is only one error, and it is from among P_1, P_2 or P_3, then one of E_1, E_2 and E_3 will be 1's.
- The ML decoder will pick the result with the fewest errors.

For $E_1 E_2 E_3 = 000$, the result is no errors.

Problem 8

Solution. Main points to consider:

- no errors result in 0 for all of E_1, E_2 and E_3.
- if there is only one error, and it is from among D_1, D_2 or D_3, then two of E_1, E_2 and E_3 will be 1's.
- if there is only one error, and it is from among P_1, P_2 or P_3, then one of E_1, E_2 and E_3 will be 1's.
- The ML decoder will pick the result with the fewest errors.

For $E_1 E_2 E_3 = 000$, the result is no errors.

For $E_1 E_2 E_3 = 010$, the result is 1 error in P_2.

Problem 8

Solution. Main points to consider:

- no errors result in 0 for all of E_1, E_2 and E_3.
- if there is only one error, and it is from among D_1, D_2 or D_3, then two of E_1, E_2 and E_3 will be 1's.
- if there is only one error, and it is from among P_1, P_2 or P_3, then one of E_1, E_2 and E_3 will be 1's.
- The ML decoder will pick the result with the fewest errors.

For $E_1E_2E_3 = 000$, the result is no errors.

For $E_1E_2E_3 = 010$, the result is 1 error in P_2.

For $E_1E_2E_3 = 101$, the result is 1 error in D_1.

Problem 8

Solution. Main points to consider:

- no errors result in 0 for all of E_1, E_2 and E_3.
- if there is only one error, and it is from among D_1, D_2 or D_3, then two of E_1, E_2 and E_3 will be 1's.
- if there is only one error, and it is from among P_1, P_2 or P_3, then one of E_1, E_2 and E_3 will be 1's.
- The ML decoder will pick the result with the fewest errors.

For $E_1E_2E_3 = 000$, the result is no errors.

For $E_1E_2E_3 = 010$, the result is 1 error in P_2.

For $E_1E_2E_3 = 101$, the result is 1 error in D_1.

For $E_1E_2E_3 = 111$, the result is multiple errors.

Theoretical bounds

Singleton bound: for any C(n, k) code,

$$d_{\text{min}} \leq n - k + 1.$$

If there is equality in the Singleton bound, we say that the code is Maximum Distance Separable (MDS).

Theoretical bounds

Singleton bound: for any $C(n, k)$ code,

$$d_{\min} \leq n - k + 1.$$

If there is equality in the Singleton bound, we say that the code is Maximum Distance Separable (MDS).

Hamming bound: if a $C(n, k)$ binary code can correct t errors, then

$$\sum_{i=0}^{t} \binom{n}{i} \leq 2^{n-k}.$$

If there is equality in the Hamming bound, we say that the code is perfect.

Problem 9

Is the following code an MDS (Maximum Distance Separable) code?

$$c_1 = (00000), \qquad c_2 = (11111)$$

Problem 9

Is the following code an MDS (Maximum Distance Separable) code?

$$c_1 = (00000), \qquad c_2 = (11111)$$

Solution. $n = 5$ $2^k = 2$, so $k = 1$, and $d_{\text{min}} = 5$. A code is MDS if

$$d_{\text{min}} = n - k + 1$$

holds. In this case,

$$d_{\text{min}} = 5 = n - k + 1 = 5 - 1 + 1$$

holds, so this is an MDS.

Problem 10

Is the following code an MDS code?

$$\begin{array}{cccc} (000000), & (001011), & (010110), & (011101) \\ (100101), & (101110), & (110011), & (111000) \end{array}$$

Problem 10

Is the following code an MDS code?

$$\begin{array}{cccc} (000000), & (001011), & (010110), & (011101) \\ (100101), & (101110), & (110011), & (111000) \end{array}$$

Solution. $n = 6, k = 3, d_{\text{min}} = 3$.

$$n - k + 1 = 6 - 3 + 1 \neq d_{\text{min}} = 3$$

No, it is not MDS.

Problem 11

A perfect code with $n = 15$ corrects $t = 1$ error. What is the value of k?

Problem 11

A perfect code with $n = 15$ corrects $t = 1$ error. What is the value of k?

Solution. For a perfect code, there is equality in the Hamming bound:

$$\sum_{i=0}^{t} \binom{n}{i} = 2^{n-k}$$

$$n + 1 = 2^{n-k}$$

$$16 = 2^{15-k},$$

so $k = 11$.